ORIGINAL ARTICLE

bioethics WILEY

Artificial intelligence and the doctor–patient relationship expanding the paradigm of shared decision making

Giorgia Lorenzini[1] | Laura Arbelaez Ossa[1] | David Martin Shaw[1,2] | Bernice Simone Elger[1,3]

[1]Institute for Biomedical Ethics, University of Basel, Basel, Switzerland

[2]Care and Public Health Research Institute, Universiteit Maastricht, Maastricht, The Netherlands

[3]Center for Legal Medicine, Unit for Health Law and Humanitarian Medicine, University of Geneva, Geneva, Switzerland

Correspondence
Giorgia Lorenzini, Institute for Biomedical Ethics, University of Basel, Bernoullistrasse 28, Basel 4056, Switzerland.
Email: giorgia.lorenzini@unibas.ch

Abstract

Artificial intelligence (AI) based clinical decision support systems (CDSS) are becoming ever more widespread in healthcare and could play an important role in diagnostic and treatment processes. For this reason, AI-based CDSS has an impact on the doctor–patient relationship, shaping their decisions with its suggestions. We may be on the verge of a paradigm shift, where the doctor–patient relationship is no longer a dual relationship, but a triad. This paper analyses the role of AI-based CDSS for shared decision-making to better comprehend its promises and associated ethical issues. Moreover, it investigates how certain AI implementations may instead foster the inappropriate paradigm of paternalism. Understanding how AI relates to doctors and influences doctor–patient communication is essential to promote more ethical medical practice. Both doctors' and patients' autonomy need to be considered in the light of AI.

KEYWORDS
artificial intelligence, autonomy, doctor–patient relationship, healthcare, shared decision-making

1 | INTRODUCTION

The introduction of *artificial intelligence* (AI) technologies in healthcare promises enormous benefits, particularly concerning quality, costs, efficiency, and access. AI applications are broad and diverse: the present analysis considers the implications of AI used in hospitals as clinical decision support systems (CDSS). CDSS can suggest diagnoses, make predictions, and recommend treatments, thus assisting physicians, nurses, patients, and other caregivers' decision-making processes.[1] CDSSs are not a novelty: computer-based ones already existed in the 1970s.[2] However, they were poorly integrated with patient care and it was not until recently, combined with AI and electronic health records (EHR), that they started to become increasingly desirable for clinical practice. Thanks to the embedding of AI and EHR with CDSS, they can provide valuable diagnostic suggestions based on patient data and test results, as well as support patient safety, clinical management, cost containment, and administrative functions.[3] AI-based CDSS are gaining momentum as they promise faster and more accurate

---

[1]Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *Npj Digital Medicine*, *3*(1), 1–10. https://doi.org/10.1038/s41746-020-0221-y;

Sloane, E. B., & Silva, R. J. (2020). Artificial intelligence in medical devices and clinical decision support systems. In E. Iadanza (Ed.), *Clinical engineering handbook* (pp. 556–568). https://doi.org/10.1016/B978-0-12-813467-2.00084-5
[2]Sutton, R. T., et al., op. cit. note 1.
[3]Ibid.

---

decisions and diagnoses.[4] As the focus of this paper lies on CDSS applications of AI, the general term 'AI' will be used as an abbreviation of AI-based CDSS.

Any tool aiming to enhance doctors' diagnostic abilities and quality of care may be life-saving for many people: diagnostic errors alone contribute to approximately 10% of patients' death in the United States.[5] However, the adoption of these technologies is not unproblematic. AI generates new questions and challenges for the doctor–patient relationship as it bears the potential to transform clinical interaction modes[6]: 'although AI systems have the potential to empower humans in medical decision-making, they also run the risk of limiting autonomy and creating new obligations'.[7] This paper focuses on the impact that AI can have on the doctor–patient relationship while trying to identify its benefits and burdens for the shared decision-making (SDM) paradigm. Although the consequences of AI for SDM are beginning to be discussed by academics (e.g., Eric Topol's *Deep Medicine*[8]), this paper aims to contribute to the discussion by providing a cautionary tale of the potential consequences of AI for both doctors and patients. After introducing the SDM paradigm, it will show how AI could relate to it, and evaluate the importance of both doctors' and patients' understanding, communication, and autonomy. There is a danger of a shift back towards paternalism if sufficient attention is not given to preserving the foundations of SDM, namely doctors' and patients' understanding, communication, and autonomy.

2 | THE SHARED DECISION-MAKING PARADIGM

Nowadays, much attention is paid to patients' autonomy, and it is generally thought that doctors should facilitate patients' participation in managing their health. Patients' right to direct their own care, namely to hold views, make choices, and act according to their personal values, should be acknowledged.[9] Respect for patients' autonomy is considered to be one of the fundamental principles of contemporary medical practice: 'respect for autonomy is not a mere *ideal* in health care; it is a professional *obligation*. The autonomous choice is a *right*—not a *duty*—of patients'.[10] Therefore, it can be argued that a paradigm promoting patients' autonomy is more ethical than one suppressing it. Indeed, it is broadly accepted as the ethically appropriate paradigm.[11] The SDM paradigm can empower patients and get them more involved in their healthcare, hence allowing them to exercise their values and autonomy.[12] By promoting SDM, patients' self-determination is promoted too. Ideally, this collaborative doctor–patient relationship should be an encounter between two experts: physicians are experts in medicine while patients are experts in their own values.[13] Both sides need to respect the other's expertise, and have the duty to inform the other: the doctors should disclose the procedure and associated risks and benefices, possible alternatives, prognosis, and consequences of each clinical decision, and the patients should indicate their preferences and personal values. Doctors no longer 'care for' as much as 'care with' their patients.[14] SDM is a process and, as such, it involves many factors that can either contribute to or hinder the shared decision. An element that can foster this process is informed consent since it calls for doctors to disclose and explain information to patients. Therefore, informed consent can be a part of SDM: without adequate information, patients have an inadequate basis for decision-making.[15] Informed consent is the disclosure of appropriate information to competent patients who then can actively participate in decisions concerning their health.[16] The information given to patients forms the basis for them to exercise their autonomy: without this information, patients would not be able to consciously contribute to decision-making. Only when patients have received enough information SDM can take place.

When AI is included in the relationship, it can support SDM, if carefully implemented. This is because AI plays a role in the decision-making process. The aspects that need to be considered are how AI influences doctors' and patients' communication and autonomy. In the event that CDSS is used without careful attention to these aspects, it could lead to a paternalistic doctor-patient relationship. While SDM has been identified as the ethically appropriate paradigm for healthcare, paternalism disregards patients' values, understanding, and autonomy.

According to the paternalistic paradigm, doctors attempt to steer patients' decisions to what they think is in the patients' best interest, for example, by strongly recommending a course of action or by offering them a partial range of options. In the extreme form of paternalism, patients may have no decisional power, but that is rarely

[4]Wang, D., Wang, L., Zhang, Z., Wang, D., Zhu, H., Gao, Y., Fan, X., & Tian, F. (2021). "Brilliant AI doctor" in rural China: Tensions and challenges in AI-powered CDSS deployment. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 6, 1–18. https://doi.org/10.1145/3411764.3445432

[5]Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, & The National Academies of Sciences, Engineering, and Medicine. (2015). Summary. In E. P. Balogh, B. T. Miller, & J. R. Ball (Eds.), *Improving diagnosis in health care* (pp. 1–18). National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK338596/

[6]Braun, M., Hummel, P., Beck, S., & Dabrock, P. (2021). Primer on an ethics of AI-based decision support systems in the clinic. *Journal of Medical Ethics*, 47(12), e3. https://doi.org/10.1136/medethics-2019-105860

[7]Rajpurkar. P., Chen, E., Banrjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28(1), 31–38. https://doi.org/10.1038/s41591-021-01614-0

[8]Eric, J., & Topol, E. J. (2019). *Deep medicine: How artificial intelligence can make healthcare human again* (1st ed.). Basic Books.

[9]Jauhar, S. (2014, February 22). Opinion | When doctors need to lie. *The New York Times*, sec. Opinion. https://www.nytimes.com/2014/02/23/opinion/sunday/when-doctors-need-to-lie.html; Beauchamp, T. L., & Childress, J. F. (1979). *Principles of biomedical ethics*. Oxford University Press.

[10]Beauchamp & Childress, op. cit. note 9.

[11]McDougall, R. J. (2019). Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*, 45, 156–160.

[12]Stiggelbout, A. M., Van der Weijden, T., De Wit, M. P. T., Frosch, D., Légaré, F., Montori, V. M., Trevena, L., & Elwyn, G. (2012). Shared decision making: Really putting patients at the centre of healthcare. *BMJ (Clinical Research Ed.)*, 344, e256. https://doi.org/10.1136/bmj.e256

[13]Bordin, E. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, and Practice*, 16(3), 252–260. https://doi.org/10.1037/H0085885; Godolphin, W. (2009). Shared decision-making. *Healthcare Quarterly*, 12(Spec No Patient), e186–e190. https://doi.org/10.12927/hcq.2009.20947

[14]Jauhar, op. cit. note 9.

[15]Beauchamp & Childress, op. cit. note 9.

[16]Appelbaum, P. S. (2007). Clinical practice. Assessment of patients' competence to consent to treatment. *The New England Journal of Medicine*, 357(18), 1834–1840. https://doi.org/10.1056/NEJMcp074045

the case: usually, paternalistic doctors try to convince patients to choose whichever option think is best. When decisions have to be made, paternalistic doctors may override patients' wishes.[17] The doctors' decisions are not always in line with the patients' wishes, which often remain unheard of. Indeed, in a paternalistic doctor–patient relationship, the doctors act on the patient's behalf, but not at the patient's behest.[18] This does not mean that doctors intend to mistreat patients, rather, their nonobservance of the patients' autonomy has the ultimate goal of doing what they consider to be the best for them. However, without opening a dialogue with patients, doctors cannot know if what they consider to be best is truly good from the patient's point of view. In this situation, it is difficult for patients to understand what is happening to their bodies, because they are not given sufficient explanations about their diagnosis, prognosis, and treatment options, with the correlated benefits and risks. Therefore, a paternalistic doctor would not seek the patient's consent and might not disclose all relevant information. Today's doctors who tend towards paternalistic behaviour, although under the legal obligation to ask for patients' consent, would try to limit information and influence patients towards what they believe is best for them. Thus, paternalistic doctors do not promote their patients' autonomy.

This paper holds SDM as the opportune paradigm for the doctor-patient relationship. Instead, a paternalistic relationship is considered undesirable in principle. Accordingly, it conceptualises different ways in which AI-CDSS could preserve SDM while trying to tackle those aspects that may promote paternalism.

3 | ARTIFICIAL INTELLIGENCE AND SHARED DECISION-MAKING

The doctor–patient relationship is a consensual partnership in which patients seek and accept the assistance of a physician to manage their health. They collaborate to achieve the highest standard of care while respecting patients' autonomy, communicating and explaining options, and obtaining informed consent.[19] Therefore, the key elements of the relationship between doctors and patients are effective communication and respect for voluntary choices.[20] These are the preconditions for SDM, and their compliance is independent of the presence of AI. However, with the introduction of AI into the equation, the enforcement of these key elements may be at risk: both the interaction between doctors and AI and the communication of this interaction to patients should be considered.

3.1 | AI–doctor communication and autonomy

During the clinical evaluation process, doctors assess complex clinical evidence to reach a diagnosis. When AI is used to assist in diagnosis, its suggestions will become part of this evaluation process. AI's suggestions could guide doctors' decisions more than they are aware of since their outputs affect, shape, and even stand in tension with doctors' judgements, thus raising questions on who is truly guiding the decision-making process.[21] AI's extensive influence could limit doctors' autonomy.[22] Accordingly, some doctors worry that AI could 'decrease their control over decision making' and that it may be a 'threat to professional autonomy'.[23] Doctors' professional autonomy is a condition for freely exercising their clinical judgement in patient care: doctors need to have the necessary autonomy to take decisions on the care of their patients.[24] This autonomy also corresponds to a responsibility on the part of doctors to provide their best care to patients.[25] The close link between doctors' autonomy and responsibility is due to the nature of this autonomy: it is granted to doctors because doing so provides a benefit to society, and that benefit would be good care.[26] Therefore, doctors' professional autonomy is important because it is a requisite for practising their judgement and decision-making, while also holding them account-able for those, and it needs to be preserved as it provides a service to society. What is at stake here is doctors' freedom to decide both the conditions for practice—for example, how AI will be implemented—and to act according to their best clinical judgement 'to promote patients' best interest, not their best interest'.[27] Since SDM is a collaborative relationship, it is essential that both parties preserve their ability to make informed and autonomous decisions. Indeed, doctors are an active part of the SDM process and, as such, they shall be capable of autonomous clinical judgement and decision-making. This requires both doctors' and patients' autonomy to be respected. Two requirements for doctors' autonomy are competence, both clinically and with AI, and ability to make their own decisions based on clinical and contextual evidence.

Competent doctors collaborate with AI while assessing its recommendations and checking for errors.[28] Of foremost importance is to identify which level of understanding is necessary for doctors to integrate AI's recommendations into daily practice while maintaining a critical eye. If doctors understand the implications and underlying assumptions of AI, they will be better positioned to evaluate its

[17]McKinstry, B. (1992). Paternalism and the doctor-patient relationship in general practice. *The British Journal of General Practice*, 42(361), 340–342.
[18]Jauhar, op. cit. note 9.
[19]Ha, J. F., & Longnecker, N. (2010). Doctor-patient communication: A review. *The Ochsner Journal*, 10(1), 38–43.
[20]Chandra, S., Mohammadnezhad, M., & Ward, P. (2018). Trust and communication in a doctor- patient relationship: A literature review. *Journal of Healthcare Communications*, 3(3), 1–6. https://doi.org/10.4172/2472-1654.100146

[21]Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751–752. https://doi.org/10.1126/science.aat5991; Braun, M., et al., op. cit. note 6.
[22]Rajpurkar, P., et al., op. cit. note 7; Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical decision support in the era of artificial intelligence. *JAMA*, 320(21), 2199. https://doi.org/10.1001/jama.2018.17163
[23]Wang, D., et al., op. cit. note 4.
[24]Wilson, C. B. (2013, March 25). *Physician autonomy essential to patient care*. Retrieved October 19, 2022, from https://www.wma.net/blog-post/physician-autonomy-essential-to-patient-care/
[25]Ibid.
[26]McAndrew, S. (2019). Internal morality of medicine and physician autonomy. *Journal of Medical Ethics*, 45(3), 198–203. https://doi.org/10.1136/medethics-2018-105069
[27]Emanuel, E. J., & Pearson, S. D. (2012). Physician autonomy and health care reform. *JAMA*, 307(4), 367–368. https://doi.org/10.1001/jama.2012.19
[28]Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, 46, 205–2011; Rajpurkar, P., et al., op. cit. note 7; Shortliffe & Sepúlveda, op. cit. note 22.

outputs, thus more confidently deciding whether or not to rely on them for their own decision-making. Hence, physicians may seek to understand the reasons underlying AI's recommendations to evaluate their validity and to be able to explain to patients their impact on the clinical evaluation process.[29] One obstacle to doctors' understanding can be the opaqueness of some AI systems, namely the *black box* problem. Black box AI increases the complexity of the communication process as it does not offer explanations of its decisions and operations.[30] Solving the black box problem can support the AI–doctor relationship, but requirements for AI to be explainable should be endorsed only if explanations consider the specific context, background knowledge, and interests of doctors, rather than being solely mathematical.[31] In the explanatory process, several factors should be included, such as premises, implications, and the AI's output in relation to the real-life context.[32] However, it must be noted that this should not be regarded as a *deus ex machina* solution. The offer of causal explanations of AI behaviour and the apparent transparency of knowing the causal relationships between the input and output does not necessarily translate to understanding the implications of using AI and its assumptions. In the same way, understanding AI's causal inferences may not be always required to evaluate its recommendation. It is above all fundamental for doctors to be aware of AI's usability and limitations in the context of implementation. Therefore, doctors would not need to know everything about AI and how it arrived at a certain recommendation, but rather, they would benefit from understanding the underlying assumptions of a decision. For example, is AI basing its analysis on similar data/situations or is it considering family history to increase or decrease a risk assessment?

While explainability may be useful, excessively concentrating on the black box issue can potentially overshadow other issues that can equally or more strongly impede doctors' capacity to work with AI. Supposing that there is a fully transparent and highly performing AI system, the problem remains as to how to train doctors to understand and evaluate its results to a degree they are competent enough to remain autonomous and decide how, when, and if to integrate AI in their clinical judgement. Explainability alone does not guarantee AI–doctor or doctor–patient communication; rather, motivation and time constraints may be equally important factors to be addressed.

Another consideration in terms of optimal AI–doctor collaboration and preserving physicians' professional autonomy, is the assistive nature of these tools: they are designed to inform, assist, and empower clinicians, not to replace them.[33] It is unlikely that in the near future, AI will replace humans as the final decision-makers in the healthcare context.[34] Despite this, there is a strong narrative suggesting that AI is a competitor with doctors, standing against their expertise, and undermining their profession. For example, a venture capitalist from Silicon Valley once proclaimed that 'machines will replace 80% of doctors' and 'radiologists will be obsolete in five years'.[35] Although they made these statements a long time ago, this has not occurred; currently, AI can only be a tool for clinicians and not a substitution.[36] Describing AI as a rival and autonomous agent does not foster a good ground for introducing AI as a further collaborator in clinical practice.

Ultimately, current AI-based tools lack the contextual and emotional intelligence needed to make decisions in uncertain, risky, and emotionally fraught circumstances: 'some decisions are not simply a matter of survival-based logic'.[37] The conclusion is not to avoid using AI until when they are 'intelligent' enough to be autonomous deciders. On the one hand, this is something that we may never want to happen; on the other hand, the underuse of AI could increase the risk of harm to patients and be burdensome.[38] If it is not desirable for doctors to avoid using AI, neither is it desirable to exclude doctors from the clinical decision-making process: clinical practice is more likely to implement human-in-the-loop setups, where doctors actively collaborate with AI systems, provide oversight, and decide what, when, how, and why to integrate its outputs in their clinical judgement.[39] What should be advocated is then a collaborative partnership between AI and doctors: AI systems should collaborate with humans instead of competing against them.[40] This collaboration would allow doctors to exercise their autonomy and to take care of those aspects that even a perfect algorithm cannot handle (such as empathy, risk communication, and assessment of patients' values, hopes, fears, and expectations[41]) while also promising better performance.[42] By joining forces, AI and doctors may provide better care than either AI or clinicians alone.[43] This is crucial because doctors' professional autonomy is a

[29]Diprose, W. K., Buist, N., Hua, N., Thurier, Q., Shand, G., & Robinson, R. (2020). Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. *Journal of the American Medical Informatics Association: JAMIA, 27*(4), 592–600. https://doi.org/10.1093/jamia/ocz229

[30]Grote & Berens, op. cit. note 28.

[31]Páez, A. (2019). The pragmatic turn in explainable artificial intelligence (XAI). *Minds and Machines, 29*(3), 441–459. https://doi.org/10.1007/s11023-019-09502-w

[32]Ibid.; Ossa, L. A., Starke, G., Lorenzini, G., Vogt, J. E., Shaw, D. M., & Elger, B. S. (2022). Re-focusing explainability in medicine. *Digital Health, 8*, 20552076221074488. https://doi.org/10.1177/20552076221074488

[33]Shortliffe & Sepúlveda, op. cit. note 22; Sloane & J. Silva, op. cit. note 1.

[34]Birch, J., Creel, K. A., Jha, A. K., & Plutynski, A. (2022). Clinical decisions using ai must consider patient values. *Nature Medicine, 28*, 229–232. https://doi.org/10.1038/s41591-021-01624-y

[35]Farr, C. (2017, April 7). Here's why one tech investor thinks some doctors will be "obsolete" in five years. *CNBC.* https://www.cnbc.com/2017/04/07/vinod-khosla-radiologists-obsolete-five-years.html

[36]Krittanawong, C. (2018). The rise of artificial intelligence and the uncertain future for physicians. *European Journal of Internal Medicine, 48*, e13–e14. https://doi.org/10.1016/j.ejim.2017.06.017

[37]Liu, X., Keane, P. A., & Denniston, A. K. (2018). Time to regenerate: The doctor in the age of artificial intelligence. *Journal of the Royal Society of Medicine, 111*(4), 113–116. https://doi.org/10.1177/0141076818762648

[38]Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2021). An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. In M. Taddeo (Ed.), *Ethics, governance, and policies in artificial intelligence* (Vol. 144, pp. 19–39). Springer. https://doi.org/10.1007/978-3-030-81907-1_3

[39]Rajpurkar, P., et al., op. cit. note 7.

[40]Ibid.

[41]Liu, X., et al., op. cit. note 37.

[42]Rajpurkar, P., et al., op. cit. note 7.

[43]Patel, B. N., Rosenberg, L., Willcox, G., Baltaxe, D., Lyons, M., Irvin, J., Rajpurkar, P., Amrhein, T., Gupta, R., Halabi, S., Langlotz, C., Lo, E., Mammarappallil, J., Mariano, A. J., Riley, G., Seekins, J., Shen. L., Zucker, E., & Lungren, M. (2019). Human–machine partnership with artificial intelligence for chest radiograph diagnosis. *Npj Digital Medicine, 2*(1), 1–10. https://doi.org/10.1038/s41746-019-0189-7

prerequisite for SDM with AI: only if doctors' autonomy is preserved, they can promote patients' autonomy as this allows them to communicate transparently with patients and explain difficult information.[44] As a consequence, patients are better positioned to participate in the SDM process even if AI is used. Otherwise, a double paternalism[45] would be established: first between AI and doctors, with the latter doing as they are told, then between doctors and patients, with patients doing as they are told in turn and the doctors left as intermediaries. In this scenario, it would be more difficult for doctors to consider results, detect errors, and disagree with a paternalistic AI. AI has the potential to sustain doctors' autonomy, but only on the precondition of good communication; otherwise, they may become mere passive executors of AI's decisions. Indeed, some experts fear that doctors may become less the deciders, and more the messengers of AI's outputs.[46] AI–doctor communication enables doctors to actively participate in the decision-making process[47] because this comprehension makes them more aware of the motives of their (dis)agreement. A doctor's ability to make decisions autonomously is an essential part of the SDM process. Accordingly, good AI–doctor communication is essential both for doctors' autonomy and for SDM.

3.2 | Doctor–patient communication and autonomy with AI

AI can have an enormous role in shaping doctors' decisions, so doctors may be required to inform their patients when AI is included in the clinical evaluation.[48] Providing this type of information to patients may help them understand better the reasons for a diagnosis, the different alternatives, and the prognosis. As a consequence, patients would be better positioned to participate in the decision-making process. While explainability can contribute to doctors' understanding and evaluation of AI's recommendations, alone, it is not sufficient to safeguard patients' autonomy, as has been previously argued. Doctors not only need to assess AI's suggestions but also need to be able and willing to communicate with patients. At the core of doctor–patient collaboration, there is the willingness of both parties to communicate. AI alone cannot ensure that this communication takes place. AI is not a threat to patients' autonomy only if doctors are predisposed to disclose and discuss it. The prerequisite for this is good communication between AI and doctors. Therefore, AI–doctor

communication not only serves the function of preserving doctors' autonomy but also enables doctors to include patients in this decision-making process, thus fostering patients' autonomy.

AI poses a risk of establishing a new form of paternalism, where the 'computer knows best'.[49] This is a possibility because AI's recommendations might not take into consideration patients' values; for example, the only value guiding its recommendation might be the goal of maximising lifespan. While it can be argued that it is a common shared value, it is also true that not all patients aim to prolong their lives: at the same stage of a terminal disease, one patient will choose palliation, while another will opt for further therapy.[50] 'In clinical settings, there can be no one-size-fits-all decision threshold'.[51] Medical decisions are not based solely on clinical information but are intertwined with preferences, values, risk tolerance, and many other personal factors that must be weighed in the decision-making process.[52] AI can better support patients by considering their preferences and unique situations: while the doctors are important intermediaries and can enquire about patients' preferences and ensure these are considered, patients would have one more guarantee that their values were being respected if these preferences were already included in the AI evaluation.

Currently, the values behind AI decisions are hidden behind the algorithm; moreover, companies and institutions, rather than patients, influence these values.[53] The first required step is to identify and expose the values embedded in the AI. Therefore, doctors (and possibly patients) should be aware of AI core values, and ensure that, eventually, patients' values are safeguarded and prevail over competing views. Doctors should ensure that patients' specific preferences are taken into account, thus facilitating SDM between patients, doctors, and AI. It could be imagined to incorporate patients' preferences and risk-taking attitudes in the algorithm so that they would be considered, for example, when proposing a treatment: 'respect for patients' autonomy means that patients' values should drive the ranking process'.[54] That would be the second step (together with a successful AI-doctor-patient communication) to attaining an optimal AI–doctor–patient partnership. Ensuring that AI respects patients' autonomy is fundamental for avoiding paternalism and enabling patients to participate in SDM.

4 | CONCLUSION

It is certainly challenging to introduce AI in healthcare, but this should not be a sufficient reason to desist. While AI may profoundly alter the doctor–patient relationship, this change is not necessarily for the worst since it could further foster SDM, if carefully implemented. However, it should be borne in mind that this will also involve a paradigm shift: while

[44]Godolphin, op. cit. note 13.

[45]The concept of "double paternalism" is not new, nonetheless, in this context it is used in a new way. Traditionally, double paternalism refers to a combination of medical and social paternalism in the case involuntary hospitalisations, for example, forced psychotherapy. However, here double paternalism means the establishment of a paternalistic relationship first between the AI and the doctors, and consequently between the doctors and the patients. These two usages of the "double paternalism" concept are therefore different.

[46]Cohen, I. G., & Graver, H. (2019, August 4). A doctor's touch: What big data in health care can teach us about predictive policing (SSRN Scholarly Paper). Social Science Research Network. https://doi.org/10.2139/ssrn.3432095; McDougall, op. cit. note 11.

[47]Quintarelli, S., Corea, F., Fossa, F., Loreggia, A., Sapienza, S. (2019). AI: profili etici. Una prospettiva etica sull'Intelligenza Artificiale: principi, diritti e raccomandazioni. BioLaw Journal—Rivista di BioDiritto, 18(3), 183–204. https://doi.org/10.15168/2284-4503-448

[48]Lorenzini, G., Shaw, D. M., Ossa, L. A., & Elger, B. S. (2022). Machine learning applications in healthcare and the role of informed consent: Ethical and practical considerations. Clinical Ethics, 1–6. https://doi.org/10.1177/14777509221094476

[49]McDougall, op. cit. note 11.

[50]Liu, X., et al., op. cit. note 37.

[51]Birch, J., op. cit. note 34.

[52]McDougall, op. cit. note 11.

[53]Liu, X., et al., op. cit. note 37.

[54]McDougall, op. cit. note 11.

SDM principles may not vary, the fundamental relationship that lies at their core will. Not only could the modes of interaction be altered, but the parties involved will be different as well. As AI is increasingly being implemented, the SDM dual relationship should be redrawn as a triad, involving the patient, the doctor, and the AI. The introduction of AI shifts the medical relationship paradigm to a new form of SDM that is shared between AI, doctors, and patients.

The new triadic SDM relationship should ensure good AI–doctor–patient communication. This could be attained on the hand, by ensuring that doctors have the competence to understand and evaluate AI's outputs while bearing in mind its limitations. On the other hand, patients should be informed of AI's involvement to allow them to better participate in the SDM process. Therefore, decision-making is truly shared when both doctors and patients are in a position to contribute, each with their unique expertise (would this be medical knowledge, contextual clues, empathy, or personal values and preferences), to the final decision, even when an AI is involved.

Including both doctors and patients in the AI decision-making process should guarantee that patients' values and preferences are considered, thus preserving their autonomy and that doctors' professional autonomy is safeguarded. A similar collaborative relationship allows for AI, doctors, and patients to join forces. Eventually, this collaboration could result in better care.

ORCID

Giorgia Lorenzini https://orcid.org/0000-0002-9155-4724
Laura Arbelaez Ossa https://orcid.org/0000-0002-8303-8789
David Martin Shaw https://orcid.org/0000-0001-8180-6927

AUTHOR BIOGRAPHIES

Giorgia Lorenzini is currently a PhD candidate at the Institute for Biomedical Ethics of Basel University, Switzerland. She has a background in philosophy, having studied at Catholic University in Milan, Italy, and KU Leuven, Belgium. She is interested in the ethical implications of the latest technological innovations, especially in healthcare.

Laura Arbelaez Ossa is a PhD candidate at the Institute for Biomedical Ethics Basel, University of Basel, Switzerland. She is interested in the ethical challenges of using artificial intelligence and has published about the explainability of AI in healthcare and how to educate doctors in AI. She has a background in medicine from Colombia and a master's in public health from the University of Edinburg

David Martin Shaw is a senior research fellow at the Institute for Biomedical Ethics at Basel University and an associate professor of health ethics and law at the Care and Public Health Research Institute at Maastricht University. He has a MA in philosophy and English literature from the University of Glasgow, an MSc in philosophy from the University of Edinburgh, and a PhD from the University of Lausanne. He is interested particularly in public health ethics, research ethics, and shared decision-making.

Bernice Simone Elger is an internist head of the Institute for Biomedical Ethics (University of Basel) and full professor at the Center for Legal Medicine (University of Geneva) where she leads the Unit for Health Law and Humanitarian Medicine. She studied medicine, theology, and bioethics in Germany, France, Switzerland, and the United States.